Fast Clustering of Equivalent Structures in Crystal Structure Prediction

BOUKE P. VAN EIJCK, JAN KROON

Department of Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

Received 21 May 1996; accepted 16 October 1996

ABSTRACT: Most methods of crystal structure prediction generate many trial structures. Because these may differ in choice of unit cell, it is not always immediately obvious whether or not two such structures are equivalent. A method to answer this question is described for the case where the asymmetric unit contains one molecule in a general position, defined by the rotation and translation of that molecule with respect to some reference geometry. In the comparison of two structures, the rotation needed to transform one orientation into the other is determined first. Then it is checked whether this rotation corresponds to a transformation that is compatible with the imposed space group symmetry. A final test compares the cell lengths, the cell angles, and the molecular centers of gravity after the transformation of one structure into the other. The method is implemented for triclinic, monoclinic, and orthorhombic systems and is found to be very fast in tests on hypothetical crystal structures of acetic acid. © 1997 by John Wiley & Sons, Inc. *J Comput Chem* **18:** 1036–1042, 1997

Keywords: crystal structure prediction; fast clustering; equivalent structures

Introduction

Prediction of a crystal structure without use of experimental information relevant to that particular structure is a very challenging subject.[1] During the last few years the improvement in force fields and the increase in easily available computer power have led to the appearance of quite a number of articles where such attempts are reported.[2–14] All results indicate that the number of possible solutions is large, sometimes over 100 within an energy range of 20 kJ/mol. Limiting this number to just a few highly possible candidates requires a very sophisticated treatment.

In the present article we address a more technical consequence of this unexpected profusion of possible crystal structures. Especially in early stages of the calculation where it is computationally too expensive to perform extensive energy minimizations with a complex force field, the number of possible candidates may not be numbered in hundreds but in thousands or even tens

Correspondence to: B. P. van Eijck; e-mail: vaneyck@chem.ruu.n1

CCC 0192-8651 / 97 / 081036-07

of thousands. In most search strategies many structures will be found repeatedly, possibly with different choices for the unit cell vectors. It is vital for the subsequent refinement that such equivalent solutions must be eliminated as soon as possible.

Karfunkel et al.[15] described an algorithm for this purpose. It is based on the comparison of powder diagrams and is thus entirely insensitive to any problem connected with the choice of unit cell. Elegant as this method is in principle, the computational demands are too high when the number of structures becomes really large.

A different method was proposed by Dzyabchenko.[16] His procedure is based on a minimum number of parameters: the six lattice constants and, for every distinct structural fragment, the three Eulerian angles and the center of mass displacement that define the transformation from some standard geometry to the position in the crystal. Basically, for every comparison all possible transformations allowed by space group symmetry are investigated. For triclinic and monoclinic space groups this number is in principle unlimited, so there must be the difficult decision of where to stop. Also, quite a few matrix operations are involved in every comparison.

Here we describe a simplification of such a procedure. Rather than trying all possible transformations allowed by space group symmetry, we first determine which one should apply and then check whether or not it is allowed. The method is only worked out for the case where one rigid molecule is present in a general position in the asymmetric unit. Such a structure already poses a great challenge to theoretical prediction; however, generalization to the more general case dealt with by Dzyabchenko[16] should not be too difficult except for packings of atoms or linear molecules where our method cannot be applied.

Preliminaries

DEFINITIONS

The components of the unit cell vectors **a**, **b**, and **c** in a Cartesian axes system can be written as the matrix

$$\mathbf{A} \equiv \begin{pmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{pmatrix}. \tag{1}$$

The orientation of the axes system can be freely chosen; we shall take

$$a_y = a_z = b_z = 0, \tag{2}$$

which makes **A** upper triangular. We shall use right-handed coordinate systems and allow only positive values for the diagonal elements.

It is customary to express a positional vector **r** in fractional coordinates denoted by x, y, and z:

$$\mathbf{r} = x\mathbf{a} + y\mathbf{b} + z\mathbf{c}. \tag{3}$$

This choice of notation is unfortunate because it strongly suggests that x, y, and z are expressed in a Cartesian axes system, which is not the case. Nevertheless, we shall adhere to it and describe the transformation between Cartesian and fractional coordinates by the matrix equation $\mathbf{r} = \mathbf{A}\mathbf{x}$.

The ith symmetry operation transforms a vector **x** in the asymmetric unit into an equivalent position $\mathbf{x}_i$, following

$$\mathbf{x}_i = \mathbf{K}_i\mathbf{x} + \mathbf{s}_i, \tag{4}$$

where $\mathbf{K}_i$ is a 3×3 rotational matrix with elements 0, 1 or -1 and $\mathbf{s}_i$ is a translation vector with elements that are integers or fractions. In Cartesian coordinates this formula can be written as

$$\mathbf{r}_i = \mathbf{A}\mathbf{K}_i\mathbf{A}^{-1}\mathbf{r} + \mathbf{A}\mathbf{s}_i \equiv \mathbf{K}'_i\mathbf{r} + \mathbf{A}\mathbf{s}_i. \tag{5}$$

It is seen that $\mathbf{K}'_i$ represents a pure rotation and/or inversion in Cartesian axes, so it must be independent of **A**. It can be shown that for triclinic, monoclinic, and the three orthogonal systems we have simply $\mathbf{K}'_i = \mathbf{K}_i$.

TRANSFORMATIONS

An equivalent description of the crystal lattice can be obtained after the transformation,

$$\begin{aligned} \mathbf{a}' &= N_{11}\mathbf{a} + N_{21}\mathbf{b} + N_{31}\mathbf{c}, \\ \mathbf{b}' &= N_{12}\mathbf{a} + N_{22}\mathbf{b} + N_{32}\mathbf{c}, \\ \mathbf{c}' &= N_{13}\mathbf{a} + N_{23}\mathbf{b} + N_{33}\mathbf{c}; \end{aligned} \tag{6}$$

or, in matrix notation,

$$\mathbf{A}' = \mathbf{A}\mathbf{N}. \tag{7}$$

Although we generally conform to Dzyabchenko's notation,[16] here we deviate and replace his $\mathbf{S}^T$ by **N**. The latter symbol conveys more clearly that the elements of this matrix must be integers and eliminates many transpose superscripts T.

The new matrix $\mathbf{A}'$ does not necessarily obey eq. (2), but can be made to do so by applying a pure rotation $\mathbf{W}$ to it,

$$\mathbf{A}'' = \mathbf{WA}' = \mathbf{WAN}. \tag{8}$$

The elements of $\mathbf{A}''$ can be found from the elements of $\mathbf{A}'$ because the two unit cells have the same cell edges and cell angles:

$$\begin{aligned} a''_x &= a', \\ b''_x &= b' \cos \gamma', \\ b''_y &= b' \sin \gamma', \\ c''_x &= c' \cos \beta', \\ c''_y &= c'(\cos \alpha' - \cos \beta' \cos \gamma')/\sin \gamma', \\ c''_z &= \sqrt{(c')^2 - (c'_x)^2 - (c'_y)^2}. \end{aligned} \tag{9}$$

Any positional vector $\mathbf{r}$ in the cell $\mathbf{A}$ remains the same in cell $\mathbf{A}'$, where nothing changes except the choice of axes. In the cell $\mathbf{A}''$ the vector is rotated and possibly translated,

$$\mathbf{r}'' = \mathbf{Wr} - \mathbf{A}''\boldsymbol{\tau}, \tag{10}$$

where $\boldsymbol{\tau}$ denotes the translation vector in fractional coordinates. Figure 1 illustrates the transformation.

As stated above, the transformation matrix $\mathbf{N}$ must have integer elements. Moreover, contrary to previous developments,[16,17] we do not allow left-handed coordinate systems so its determinant must be $+1$. We further require that symmetry is conserved; for instance, space group Pc should not change into $P1$ and not even into Pa or Pb. So, depending on the space group, a certain transformation $\mathbf{N}$ may or may not be allowed and each component of the translation $\boldsymbol{\tau}$ is either free to be chosen arbitrarily or fixed by symmetry. The rules to make these decisions can be found by study of the symmetry properties of the space group. For example, in ref. 7 we did so for $P2_12_12_1$. General tables to find allowed rotations and the corresponding translations were published by Koch and Fischer.[18] Incorporating these results into a computer program is a complex task, and we developed a different approach where a given transformation $\mathbf{N}$ is examined without the use of space group tables. This procedure is outlined in the next section.

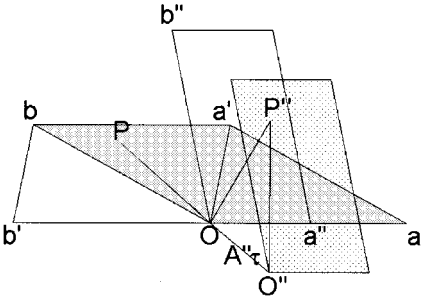

FIGURE 1. The cell $\mathbf{A}$ is transformed to $\mathbf{A}'$ ($\mathbf{a}' = \mathbf{a} + \mathbf{b}, \mathbf{b}' = -\mathbf{a}$), which is followed by a rotation to $\mathbf{A}''$ and a translation over $\mathbf{A}''\boldsymbol{\tau}$. A vector $\mathbf{r}$ (OP) is rotated over the same angle to OP″ and is represented in the new cell by $\mathbf{r}''$ (O″P″).

SYMMETRY REQUIREMENTS

The transformation to an equivalent system, defined by eq. (10), can be written in fractional coordinates,

$$\mathbf{A}''\mathbf{x}'' = \mathbf{WAx} - \mathbf{A}''\boldsymbol{\tau}; \tag{11}$$

or, with the aid of eq. (8),

$$\mathbf{x}'' = \mathbf{N}^{-1}\mathbf{x} - \boldsymbol{\tau}. \tag{12}$$

The symmetry operation i transforms $\mathbf{x}$ into $\mathbf{x}_i$ [eq. (4)], which can also be transformed into the equivalent system,

$$\mathbf{x}''_i = \mathbf{N}^{-1}(\mathbf{K}_i\mathbf{x} + \mathbf{s}_i) - \boldsymbol{\tau}. \tag{13}$$

But, if $\mathbf{x}$ is first transformed to the equivalent system and then transformed by the symmetry operation j, we find

$$x''_j = \mathbf{K}_j(\mathbf{N}^{-1}\mathbf{x} - \boldsymbol{\tau}) + \mathbf{s}_j. \tag{14}$$

Now, the two resulting structures must be the same, so to every i must correspond a certain j to make $\mathbf{x}''_i = \mathbf{x}''_j$. Comparison of eqs. (13) and (14), which must both be valid for any $\mathbf{x}$, then leads to the two equations,

$$\mathbf{K}_i\mathbf{N} = \mathbf{NK}_j \tag{15}$$

and

$$(\mathbf{N} - \mathbf{NK}_j)\boldsymbol{\tau} = \mathbf{s}_i - \mathbf{Ns}_j. \tag{16}$$

First we must check whether or not a set of combinations i, j can be found that satisfies eq. (15). In space groups with centering each $\mathbf{K}$ will occur more than once, which leads to multiple sets of combinations. For every possible set, each j corresponds to a given i and we can abbreviate the right-hand side of eq. (16) to $\mathbf{s}'_i$. Because translations along unit cell vectors are always allowed, each component of $\mathbf{s}'_i$ can be shifted by any inte-

ger. Further abbreviating the vector $\mathbf{N}\boldsymbol{\tau}$ to $\boldsymbol{\tau}'$, we reformulate the set of linear eqs. (16) (three for every symmetry operation i) as

$$(\mathbf{E} - \mathbf{K}_i)\boldsymbol{\tau}' = \mathbf{s}'_i. \quad (17)$$

It may turn out that no solutions exist, and then the transformation is not allowed after all. In the other case, possible limitations on the allowed values for the shift vector $\boldsymbol{\tau}'$ emerge. The situation can be illustrated for the simple case of triclinic, monoclinic, or orthorhombic space groups where each $\mathbf{K}_i$ is diagonal. Here we have two possibilities for every component g of every symmetry operation i: $(K_i)_{gg} = +1$: no condition for τ'_g, but $(s'_i)_g$ must be an integer; and $(K_i)_{gg} = -1$: $\tau'_g = \frac{1}{2}(s'_i)_g$ modulo $\frac{1}{2}$.

Clustering Algorithm

GENERAL METHOD

Let us suppose that we have a reference structure for a rigid free molecule in a Cartesian axes system. A vector from the center of gravity to a certain atom is defined by the vector $\mathbf{r}_0$. Let us further suppose that by any search method we obtained two possible crystal structures, numbered 1 and 2, where one such molecule forms the asymmetric unit. The corresponding vectors in these two structures can then be expressed as

$$\mathbf{r}_1 = \mathbf{R}_1\mathbf{r}_0 + \mathbf{A}_1\mathbf{t}_1,$$
$$\mathbf{r}_2 = \mathbf{R}_2\mathbf{r}_0 + \mathbf{A}_2\mathbf{t}_2, \quad (18)$$

where $\mathbf{R}_1$ and $\mathbf{R}_2$ denote pure rotational matrices and $\mathbf{t}_1$ and $\mathbf{t}_2$ are fractional translations of the center of gravity of the molecule. Together with $\mathbf{A}_1$ and $\mathbf{A}_2$, which define the unit cells, these quantities fully determine the two crystal structures. The two sets of data are the input for the following procedure that answers the question whether or not the two structures are equivalent within certain limits.

Eliminating $\mathbf{r}_0$ from eqs. (18) leads to

$$\mathbf{r}_2 = \mathbf{R}_2\mathbf{R}_1^T(\mathbf{r}_1 - \mathbf{A}_1\mathbf{t}_1) + \mathbf{A}_2\mathbf{t}_2. \quad (19)$$

Because the two structures must be related by a transformation as described above, we can compare this equation with eq. (10) if we equate structure 1 to $\mathbf{A}$ and structure 2 to $\mathbf{A}''$. We find

$$\mathbf{W} = \mathbf{R}_2\mathbf{R}_1^T, \quad (20)$$

from which we obtain the transformation matrix,

$$\mathbf{N} = (\mathbf{WA})^{-1}\mathbf{A}'' = (\mathbf{A}_1^{-1}\mathbf{R}_1)(\mathbf{R}_2^T\mathbf{A}_2). \quad (21)$$

Note that eq. (20) is different from the corresponding equation 8.1 of Dzyabchenko.[16] In that procedure the order of the matrix multiplication was interchanged for computational efficiency, which is possible because only the trace of $\mathbf{W}$ is needed there.[19]

If the two structures are equivalent, $\mathbf{N}$ must have only integer elements. However, because no search procedure can be expected to give exact results, some tolerance must be allowed here for deviations from integer values. Most pairs of structures will not be equivalent, and the test can be stopped at this point. For those that pass the test, further (exact) checks are necessary: the determinant of $\mathbf{N}$ should be $+1$, and the other symmetry requirements set in eqs. (15) and (17) must also be met.

If that is the case, the transformation from $\mathbf{A}_1$ to equivalent structures $\mathbf{A}'_1$ and $\mathbf{A}''_1$ can be performed using eqs. (7) and (9). Then all elements of $\mathbf{A}''_1$ and $\mathbf{A}_2$ should agree within a certain tolerance. Likewise, the position of the center of gravity is transformed [eq. (12)] to

$$\mathbf{t}''_1 = \mathbf{N}^{-1}\mathbf{t}_1 - \boldsymbol{\tau}. \quad (22)$$

Because $\mathbf{t}''_1$ and $\mathbf{t}_2$ should be the same, we obtain the condition,

$$\mathbf{N}\mathbf{t}_2 = \mathbf{t}_1 - \boldsymbol{\tau}', \quad (23)$$

which should be satisfied (again, within a certain tolerance) for each component of $\boldsymbol{\tau}'$ that cannot be freely chosen, as found from eqs. (17).

MOLECULES WITH INTERNAL SYMMETRY

In the previous section we tacitly assumed that the reference molecule has no internal symmetry. Let us now assume that such a symmetry operation, $\mathbf{M}$, (other than the identity) does exist: the set vectors $\mathbf{r}_0$ is equivalent to the set $\mathbf{M}\mathbf{r}_0$.

A certain crystal structure 1 is again built from the reference structure by the transformation $\mathbf{r}_1 = \mathbf{R}_1\mathbf{r}_0 + \mathbf{A}_1\mathbf{t}_1$. Equivalent to this is now the structure 1* found from

$$\mathbf{r}_1^* = \mathbf{R}_1\mathbf{M}\mathbf{r}_0 + \mathbf{A}_1\mathbf{t}_1. \quad (24)$$

If $\mathbf{M}$ represents a pure rotation, this structure might actually correspond to a structure 2 to be tested for equivalence with structure 1. However, the proce-

dure described in the previous section cannot recognize the equivalence unless an additional test is carried out. It is seen that the comparison of the set $\{\mathbf{R}_2, \mathbf{A}_2\mathbf{t}_2\}$ with the set $\{\mathbf{R}_1, \mathbf{A}_1\mathbf{t}_1\}$ has to be extended by comparison with the set $\{\mathbf{R}_1\mathbf{M}, \mathbf{A}_1\mathbf{t}_1\}$.

The situation is more complicated if $\mathbf{M}$ is a rotation-inversion operation, which includes the frequent case of a mirror plane. Obviously, structure 1* cannot be found in a search procedure that only considers proper rotations of the reference molecule. Now let us assume that sign inversion of all coordinates produces an equivalent structure (with opposite crystal chirality in enantiomorphous space groups); this is the case for all space groups of orthorhombic or lower symmetry except *Fdd*2. Then structure 1* is equivalent to a structure 1**,

$$\mathbf{r}_1^{**} = -\mathbf{R}_1\mathbf{M}\mathbf{r}_0 - \mathbf{A}_1\mathbf{t}_1, \qquad (25)$$

which can be obtained from the reference structure by a proper rotation and a translation. So this structure might be equivalent to the structure 2 to be tested, and the test for equivalence of the set $\{\mathbf{R}_2, \mathbf{A}_2\mathbf{t}_2\}$ with the set $\{\mathbf{R}_1, \mathbf{A}_1\mathbf{t}_1\}$ now has to be extended by comparison with the set $\{-\mathbf{R}_1\mathbf{M}, -\mathbf{A}_1\mathbf{t}_1\}$. These additional tests have to be performed for every independent symmetry operation $\mathbf{M}$ in the reference molecule.

Discussion

IMPLEMENTATION

The above method was implemented into a Fortran subroutine for triclinic, monoclinic, and orthorhombic systems. First, a list was made of the matrices $\mathbf{R}^T\mathbf{A}$ for all structures to be intercompared. The list was ordered after increasing energy and possibly truncated if it was excessively large, say, over 100,000 entries. In turn, each structure became a reference structure 1 for which $\mathbf{A}_1^{-1}\mathbf{R}_1$ was calculated, possibly extended with additional matrices for molecules with internal symmetry as discussed in the previous section. When N structures had to be compared, the effort for these preliminary calculations was only proportional to N. But then all structures higher in the list had to be considered as test structures 2, and the total number of comparisons to be made was proportional to N^2. For these test structures $\mathbf{R}_2^T\mathbf{A}_2$ was fetched from memory, and the calculation of $\mathbf{N}$ by eq. (21) was started. So in this time-consuming part of the calculation only one matrix multiplication had to be performed, and it could be stopped as soon as an element was encountered that deviated too much from an integer value. This occurred in the majority of cases.

The duration of the calculations was very sensitive to the amount of equivalent structures encountered because, of course, a structure was immediately removed from the list for further testing once such equivalence was established.

Two tolerances must be set: ϵ_N for the maximum allowed deviation of the elements of the $\mathbf{N}$ matrix from integer values, and ϵ_A for the difference in the cell parameters and center of gravity positions between the two structures being compared.

Results

The procedure was tested on hypothetical crystal structures of acetic acid. A brute-force grid search[7] was carried out in the six most abundant space groups for organic molecules, augmented with $Pna2_1$ which is the space group of the experimental structure. The GROMOS force field[20] was used with charges $+0.15$ e for the united CH_3 "atom," $+0.38$ e for C, -0.38 e for the carbonyl O, -0.55 e for the hydroxyl O, and $+0.40$ e for the hydroxyl H. Here we only report our experience with the clustering algorithm, starting with lists of structures obtained by the grid search where just a few cycles of energy minimization had been done.

For timing experiments, test sets (limited to 1000 structures) were studied in space groups $P\bar{1}$ and *Pbca*. Removal of equivalent structures was inhibited by setting ϵ_A to zero. With ϵ_N set to 0.2, the CPU times for these two space groups were 13 and 16 s, respectively, on a Silicon Graphics Indigo 2 workstation equipped with an MIPS R4010 processor. The difference was due to the effort of checking the symmetry of the transformation, because there were two equivalent positions in $P\bar{1}$ and eight in *Pbca*. This difference was so modest that it was not worthwhile to expend effort in replacing the symmetry tests [eqs. (15), (17)] by a possibly more efficient algorithm based on published tables for each space group. The CPU time for one comparison of two structures was of the order of 30 μs, 4000 times faster than reported by Karfunkel et al.[15] for their algorithm on a comparable machine. Dzyabchenko[16] gives no timing re-

sults; his program does not appear to be optimized for this kind of work.

A more realistic test is the actual clustering of the full sets of initial structures. In a first cycle the original lists of structures were clustered using $\epsilon_N = 0.2$ and $\epsilon_A = 1$ Å. The remaining structures were subjected to a full energy minimization, a process that is considerably more time consuming (for acetic acid, about 2 CPUs per structure in our system). Then a second cycle of clustering was performed, and again a significant reduction in the number of possible structures was obtained. The results are reported in Table I.

So the first cycle of clustering might have been done more drastically. To verify this, the process was repeated twice with larger initial values for ϵ_N and ϵ_A. A smaller set of final structures was obtained in a considerably shorter time, but at the price of a possible loss of some structures that were not equivalent after all. To obtain a quantitative indication of this loss, the low-energy solutions (within a range of 5 kJ/mol) obtained with $\epsilon_N = 0.2$ and $\epsilon_A = 1$ Å were added and the merged list was clustered again. The tolerance criteria were now chosen fairly wide, so those structures out of the added low-energy ones that were left must be considered as really not found with the larger tolerances. The numbers of these lost solutions are reported as N_{lost} in Table I. The numbers are relatively small, but they include a few structures where the energy is only 1 or 2 kJ/mol above the lowest one. As always, one has to judge whether or not a significant gain in computer time is worth taking the small risk of eliminating the correct solution in an early stage of the investigation.

We are planning to publish full details of the possible acetic acid structures in a later stage. Nevertheless, it is of interest to report here that the experimental structure was present after all clusterings performed in space group $Pna2_1$. In the ordered list of solutions of all space groups taken together its ranking is 19, with an energy of 1.6 kJ/mol above the lowest one. This is not unsatisfactory, but the differences in geometry between observed and calculated structures are uncomfortably large. At present we are trying to find a better force field to allow more reliable crystal structure predictions.

TABLE I.
Clustering Results for Acetic Acid in Seven Space Groups.

	$P\bar{1}$	$P2_1$	$P2_1/c$	$C2/c$	$P2_12_12_1$	$Pna2_1$	$Pbca$
N_0	37705	2294	60,683	69,908	7035	3874	9410
N_1 (0.2, 1.0)	5225	432	15,051	23,122	872	1111	1838
N_2 (0.2, 1.0)	810	111	4359	10,234	300	454	761
N_3	136	28	505	273	52	56	45
CPU (clus)	27.2	0.3	154.8	303.2	1.3	0.6	2.5
CPU (emin)	196.8	14.9	512.9	761.6	29.4	35.1	61.8
N_1 (0.25, 1.5)	1465	258	7799	12,742	405	768	958
N_2 (0.2, 1.0)	440	100	3066	7179	173	368	540
N_3	88	27	426	210	42	46	44
N_{lost} (0.3, 2.0)	1	1	20	6	0	4	3
CPU (clus)	5.6	0.2	71.5	155.0	0.7	0.4	1.3
CPU (emin)	53.4	9.0	265.9	421.0	13.4	23.8	30.3
N_1 (0.3, 2.0)	508	180	4598	7489	258	563	541
N_2 (0.2, 1.0)	237	79	2167	4908	151	301	373
N_3	64	21	386	233	41	43	48
N_{lost} (0.3, 2.0)	9	0	30	6	1	3	6
CPU (clus)	2.1	0.1	37.8	79.7	0.4	0.3	0.7
CPU (emin)	18.2	6.3	153.8	282.7	8.1	17.0	16.8

N_0 is the number of original structures. A first cycle of clustering leaves N_1 structures; values in parentheses denote ϵ_N and ϵ_A (Å), respectively. Energy minimization and a second cycle of clustering results in N_2 structures, of which N_3 have an energy less than 5 kJ/mol above the lowest one. After merging the latter with the first set of such structures and clustering, the increase of N_3 represents the number of lost solutions (N_{lost}). The CPU times for clustering (clus) and energy minimization (emin) are reported in minutes.

References

1. A Gavezzotti, *Acc. Chem. Res.*, **27**, 309 (1994).
2. A. Gavezzoti, *J. Am. Chem. Soc.*, **113**, 4622 (1991)
3. H. R. Karfunkel and R. J. Gdanitz, *J. Comput. Chem.*, **13**, 1171 (1992).
4. H. R. Karfunkel and F. J. J. Leusen, *Speedup*, **6**, 43 (1992).
5. H. R. Karfunkel, F. J. J. Leusen, and R. J. Gdanitz, *J. Comput.-Aided Mater. Design*, **1**, 177 (1993).
6. J. R. Holden, Z. Du, and H. L. Ammon, *J. Comput. Chem.*, **14**, 422 (1993).
7. B. P. van Eijck, W. T. M. Mooij, and J. Kroon, *Acta Crystallogr.*, **B51**, 99 (1995).
8. T. Shoda, K. Yamahara, K. Okazaki, and D. E. Williams, *J. Mol. Struct.*, **333**, 267 (1995).
9. K. D. Gibson and H. A. Scheraga, *J. Phys. Chem.*, **99**, 3765 (1995).
10. M. U. Schmidt, Ph.D. thesis, Technische Hochschule Aachen, Aachen, Germany, 1995.
11. N. Tajima, T. Tanaka, T. Arikawa, T. Sakurai, S. Teramae, and T. Hirano, *Bull. Chem. Soc. Jpn.*, **68**, 519 (1995).
12. A. Gavezotti, *Acta Crystallogr.*, **B52**, 201 (1996).
13. A. M. Chaka, R. Zaniewski, W. Youngs, C. Tessier, and G. Klopman, *Acta Crystallogr.*, **B52**, 165 (1996).
14. D. E. Williams, *Acta. Crystallogr.*, **A52**, 326 (1996).
15. H. R. Karfunkel, B. Rohde, F. J. J. Leusen, R. J. Gdanitz, and G. Rihs, *J. Comput. Chem.*, **14**, 1125 (1993).
16. A. V. Dzyabchenko, *Acta Crystallogr.*, **B50**, 414 (1994).
17. F. L. Hirschfeld, *Acta Crystallogr.*, **A24**, 301 (1968).
18. E. Koch and W. Fischer, In *International Tables for X-ray Crystallography*, vol. A, T. Hahn, Ed., Kluwer, Dordrecht, 1989, p. 855.
19. A. V. Dzyabchenko, private communication.
20. W. F. van Gunsteren and H. J. C. Berendsen, GROMOS, Groningen molecular simulation package, University of Groningen, 1987.